

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

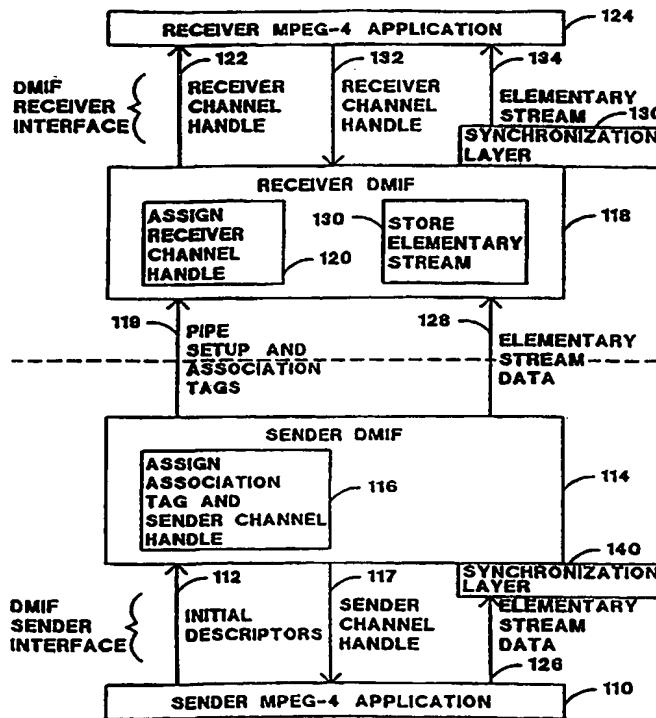
(51) International Patent Classification <sup>6</sup> : <b>H04N 7/26, 7/52, G06F 17/30</b>		A1	(11) International Publication Number: <b>WO 98/57499</b> (43) International Publication Date: 17 December 1998 (17.12.98)
(21) International Application Number: <b>PCT/IB98/01042</b> (22) International Filing Date: 8 June 1998 (08.06.98)		(81) Designated States: JP, KR, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 60/049,078 9 June 1997 (09.06.97) US 09/085,393 27 May 1998 (27.05.98) US		<b>Published</b> <i>With international search report.            Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(71) Applicant: SHARP KABUSHIKI KAISHA [JP/JP]; 22-22, Nagaike-cho, Abeno-ku, Osaka-shi, Osaka 545-0013 (JP). (72) Inventor: CRINON, Regis, J.; 2346 N.W. Cascade Street, Camas, WA 98607 (US). (74) Agent: TAKANO, Akichika; Nagasa Patent Office, Salute Building, 9th floor, 72, Yoshida-cho, Naku, Yokohama-shi, Kanagawa 231-0041 (JP).			

**BEST AVAILABLE COPY**

## (54) Title: SYSTEM FOR THE TRANSMISSION OF AUDIO, VIDEO AND AUXILIARY DATA

## (57) Abstract

A video system includes a plurality of frames of video each of which is defined by a plurality of scene elements. The scene elements for a respective frame together define an image of the frame. First auxiliary data is descriptive of a first scene element of the frame and second auxiliary data is descriptive of a second scene element of the frame. A sending device sends the frame of video, including its scene elements, the first auxiliary data, and the second auxiliary data, from the sending device to a receiving device.



***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## SYSTEM FOR THE TRANSMISSION OF AUDIO, VIDEO AND AUXILIARY DATA

The present application is a continuation of co-pending patent application, Sezen et al., Serial No. 5 60/049,078, filed June 9, 1997.

## TECHNICAL FIELD

The present invention relates to a system for encoding, sending, and receiving audio and video that 10 includes auxiliary data.

## BACKGROUND ART

Digitally encoding video, consisting of multiple sequential frames of images, is a common task 15 that traditionally has been performed by dividing each frame into a set of pixels and defining a sample luminance value, and possibly a color sample value, for each pixel. The development of powerful inexpensive computer technology has made it cost effective to develop 20 complex systems for encoding moving pictures to achieve significant data compression. This permits, for example, high definition television (HDTV) to be broadcast within a limited bandwidth. The international standard which has been chosen for the transmission of HDTV is the 25 Moving Pictures Experts Group 2 (MPEG-2) standard. The MPEG-2 video compression standard is based on both intra and inter coding of video fields or frames and achieves compression by taking advantage of either spatial or temporal redundancy in the video sequence.

An additional video encoding system has been proposed, designated as the MPEG-4 standard. MPEG-4 is a system of encoding video by which a moving picture composed of a sequence of sequential frames, generally accompanied by audio information, may be transmitted or 35 otherwise sent from a moving picture sending device ("sender") to a moving picture receiving device ("receiver") in an efficient and flexible manner.

Interactive applications are also anticipated in the MPEG-4 standard, but for ease of description, this patent application describes the case in which a moving picture is sent from the sender to the receiver only. It is to 5 be understood that this patent application includes the case in which control information may be sent from receiver to sender.

MPEG-4 supports the transmission of a "Binary Format for Scene Description" (BIFS) which specifies the 10 composition of objects in a sequence of frames or fields representing a three dimensional scene. In BIFS a scene is divided into a plurality of "scene elements." For example, as shown in FIG. 1A, a scene with a presenter, a blackboard, a desk, a globe, and an audio accompaniment 15 is broken up into scene elements in panel 12, and reconstituted into a complete scene, with all the scene elements properly arranged in panel 14. In other words, each frame of a MPEG-4 based video system is composed of 20 a plurality of scene elements that are arranged according to directives specified in the associated BIFS information. As such MPEG-4, similar to MPEG-2, is directed to a system for encoding and transmitting video and audio information for viewing a movie, albeit using a different protocol than MPEG-2. An example of a suitable systems 25 decoder for MPEG-4 is shown in FIG. 1B.

Referring to FIG. 2, each scene element 20 is represented by an object to which is assigned a node 26 in a generally hierarchical tree structure (panel 16). The scene elements requiring input streaming data include 30 an object descriptor 14. The tree structure provides a convenient data structure for representing a sequence of video fields or frames. Each object descriptor in turn includes at least one elementary stream descriptor 28 which includes such information as the data rate, 35 location information regarding the associated data stream(s) and decoding requirements for its respective logical channel (described below) for updating

information regarding the data object. The data stream sent through a particular pipe for a particular object descriptor is generally referred to as an elementary stream. Such information may include, for example, data 5 describing a change in shape, colorization, brightness, and location. Every elementary stream descriptor includes a decoder type (or equivalent stream type structure) which publishes the format or encoding algorithm used to represent the transmitted elementary 10 stream data. FIGS. 3 and 4 show a simplified object descriptor format and an elementary stream descriptor format, respectively.

In an MPEG-4 system, both the sender and the receiver can use a respective "Delivery Multimedia 15 Integration Framework" (DMIF) for assigning the data channels and time multiplexing scheme based on requests from the MPEG-4 applications. The DMIF, in effect, acts as an interface between an MPEG-4 application and the underlying transport layer, sometimes referred to herein 20 as a "data pipe." The pipes generally refer to the physical or logical interconnection between a server and a client. As an object descriptor is sent from the sender to the receiver, the sender DMIF examines each of its elementary stream descriptors 28 and assigns a pipe 25 to each one based on the requirements 30, 32, 34 and 36 included in the elementary stream descriptors 28. For example, an elementary stream with a more exacting quality of stream requirement 34 would be assigned a pipe that would permit high quality transmission. The pipe 30 assignment is sent to the receiving DMIF. For example, depending on the transmission system each pipe (logical channel) could be a different periodic time portion of a single connection; each pipe could be different connection through the internet; or each pipe could be a different telephone line. In general, DMIF establishes 35 the connections and ensures that both the sender and the receiver are in synchronization for tying a particular

pipe to a particular elementary stream at both the sending and the receiving end.

The DMIF also assigns an association tag to each elementary stream. This tag provides a single identification of each elementary stream from the sending DMIF to the receiving DMIF and vice versa. Accordingly, the association tag is a unique end-to-end DMIF pipe identifier. Any channel change or reassignment initiated by the sender or receiver (for an interactive system) would require a different association tag.

Referring to FIG. 1B, DMIF delivers data to an MPEG-4 application by way of "Access Unit Layer Protocol Data Units" (AL-PDUs). The AL-PDUs configuration descriptor 26 in the elementary stream descriptor 28 (FIG. 4) specifies the AL-PDUs time coordination and required buffer size characteristics. The purpose is to ensure that AL-PDUs will arrive in a timely manner and conveying enough data to permit the data object to be fully updated for timely display. The AL-PDUs are stored in decode buffers, are reassembled into access units, are decoded (in accordance with the Decoder Type associated with the elementary stream), and stored in a configuration buffer, to be placed into the scene by the compositor which arranges the scene elements of the scene. DMIF also configures the multiplexing and access layer for each elementary stream based on its AL-PDU Configuration Descriptor. When DMIF is not used (typically in the case of a storage network), the association tag is substituted for a channel number identifying a logical or a physical channel containing the elementary stream.

The process of sending a video stream begins with a moving picture request from the receiver application (in a broadcast system the request goes no further than the receiver DMIF), followed by a sequence of data from a sender application designed to establish an initial scene configuration. In FIG. 4, one of the

decoder types is for scene description and another is for object descriptors. In the initial scene configuration a number of object descriptors are typically established. Each data object descriptor includes at least one elementary stream descriptor.

Another issue faced in an MPEG-4 system is avoiding collision among elementary stream identifiers when two applications run simultaneously on top of the same DMIF session. Each elementary stream ID is unique within an application. Therefore, the underlying DMIF session must be able to distinguish requests from each application in the case where elementary stream ID values collide.

Hamaguchi U.S. Patent No. 5,276,805 discloses an image filing system in which retrieval data for a first set of images is associated with image data for a first image that is logically associated with the first set of images. For example, a user viewing an x-ray image of a patient's kidney would be able to view a list of a set of x-ray images taken from the same patient and be able to quickly view any of the set of x-ray images upon request. Hamaguchi teaches that the retrieval information is associated with the image as a whole.

Judson U.S. Patent No. 5,572,643 discusses an already existing technology, known as hypertext, by which an internet browser permits a user to access an internet URL by clicking on a highlighted word in a page accessed on the internet. Judson teaches that the retrieval information is associated only with the test. Cohen et al. U.S. Patent No. 5,367,621 is quite similar to Judson, permitting a user to click on a word in an on-line book, thereby causing a multimedia object to be displayed. Both Judson and Cohen et al. disclose advances linking together data which is unrelated to the encoding and data compression necessary for MPEG-2 and MPEG-4.

## DISCLOSURE OF THE INVENTION

In a first embodiment, the present invention is a video system that includes a plurality of frames of video each of which is defined by a plurality of scene elements. The scene elements for a respective frame together define an image of the frame. First auxiliary data is descriptive of a first scene element of the frame and second auxiliary data is descriptive of a second scene element of the frame. A sending device sends the frame of video, including its scene elements, the first auxiliary data, and the second auxiliary data, from the sending device to a receiving device.

Auxiliary data is data which does not directly represent audio or video but is at least one level of abstraction away from a direct audio or video representation or is information which is in some way related to the object. By including auxiliary data with the video system descriptive of the individual scene elements, in contrast to the entire frame, additional information specific to scene elements becomes available to the receiver. As an example, the additional information may be displayed on command or used to further process the video information. This information may or may not be synchronized to the data stream(s) conveying the scene element sequence.

In another embodiment of the present invention the video system includes a plurality of frames of video, each of which defines an image. Each of the frames is defined by a plurality of pixels. First auxiliary data is descriptive of a first portion of the pixels of the frame and associated with a first portion of the pixels. Second auxiliary data is descriptive of a second portion of the pixels of the frame and associated with a second portion of the pixels. A sending device sends the first portion of the pixels, the first auxiliary data, the second auxiliary data, and the second portion of the pixels to a receiving device.

The inclusion of auxiliary data into video systems that define each frame by a group of pixels, such as MPEG-2, allows sub-portions of each frame to be associated with the auxiliary data for subsequent use.

5       The foregoing and other objectives, features, and advantages of the invention will be more readily understood upon consideration of the following detailed description of the invention, taken in conjunction with the accompanying drawings.

10

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A is an illustration of the data encoding and scene reconstruction of MPEG-4.

15       FIG. 1B is an illustration of a systems decoder for MPEG 4.

FIG. 2 is an illustration of the data structure of MPEG-4.

FIG. 3 is an illustration of an MPEG-4 object descriptor.

20       FIG. 4 is an illustration of an MPEG-4 elementary stream descriptor.

FIG. 5 is an illustration of a receiver DMIF and a sender DMIF user interfaces according to the present invention.

25

#### BEST MODES FOR CARRYING OUT THE INVENTION

When MPEG-4 was first conceived, it was anticipated that it would be a system for compressing and then subsequently decompressing and displaying visual and audio data. This is in accordance with previous video compressing systems, such as MPEG-2. The present inventor came to the realization that other types of data that are specific to scene elements would be useful to the user and greatly expand the potential of an MPEG-4 system. For example, information concerning a visual object, such as the height of a building, the make or model of an automobile, a basketball player's statistics,

or a link (pointer) to related information may be of interest to a viewer. Also, statistical information about the visual object, for instance a color description describing and quantifying the coloration of a particular object may be of use to the receiver in correctly displaying the object. In addition, data describing audio elements may also be included. This type of data is, in this application, referred to as "auxiliary data." Auxiliary data is data which does not directly represent audio or video but is at least one level of abstraction away from a direct audio or video representation or is information which is in some way related to the object. Auxiliary data can be used to enhance the original video/audio service experienced by the viewers.

In a preferred embodiment of the present invention, auxiliary data may be displayed, on command, in association with a scene element. For example, a person watching a basketball game may wish to learn more about any one of the players. Television sets, or other display systems, may be equipped with a mouse, a remote command unit, or other pointing device to select scene elements being displayed. Using such a device, a viewer could point to one of the basketball players, press a button, and thereby cause the players statistics to be displayed. Such auxiliary data may be transmitted in a simple ASCII format. The user may select where the information is displayed, for example at either the top or bottom of the screen. Another use for auxiliary data is for an on-screen program guide with brief program descriptions. The brief program descriptions of the program guide provided may be selected to obtain a more complete description of the programming. Yet another use of auxiliary data is to obtain additional information about an actor while viewing a television show.

Other uses for the auxiliary data could be in conjunction with "hot buttons" displayed on the television screen. For example, there could be a hot button

for a prewritten commentary on the show, or to allow the viewer to see and to add to a running commentary by viewers watching the show.

The preferred embodiment of the present invention includes an elementary stream descriptor, similar to that shown in FIG. 2 but in which one or more decoder type (or equivalently, stream type) structures announce the presence of one or several auxiliary data streams encoded according to the format specified by decoder type. The elementary stream descriptor(s) including such decoder type (or stream type) structure may be included in an object descriptor featuring elementary stream descriptor featuring other values of decoder type (or stream type) like MPEG-4 visual streams. In this case the auxiliary data is associated with and complements the elementary streams published by these elementary stream descriptor(s). In another situation, the elementary stream descriptors announcing the presence of auxiliary data may be included in an object descriptor which does not feature elementary stream descriptors featuring other kinds of decoder type (or stream type). In this situation, the auxiliary data may be associated and may complement a scene element 20 in a node 26 which is directly linked to the node publishing the auxiliary data object descriptor.

The present invention ensures that the elementary streams and the associated auxiliary data streams can be unambiguously identified in the MPEG-4 player as a result of the DMIF assigning an association tag to each elementary stream. The association tag data is transmitted via a control message from the sender DMIF and to the receiver DMIF. A separate channel handle uniquely representative of the use of an association tag by an application is assigned at run-time by the server and the client DMIF, respectively. Not only does this provide the receiving application with a shorthand way of referencing the channel conveying the AL-PDUs, but it

avoids confusion when at least two applications using the same elementary stream identifier are running simultaneously in the same DMIF session.

FIG. 5 illustrates the assignment of association tags and channel handles. The sender MPEG-4 application 110 sends the initial object descriptors, including elementary stream descriptors (transmission 112) to the sender DMIF. The sender DMIF 114 during setup assigns an association tag and a sender MPEG application program channel handle to each requested elementary stream (step 116) and passes the sender channel handle back to the sender MPEG-4 application (step 117). The receiver DMIF 118 assigns a receiver channel handle to each requested elementary stream (step 120) and passes it (step 122) to the receiver MPEG-4 application 124, thus creating an end-to-end pipe indicator composed of the sender channel handle, association tag, and receiver channel handle. When elementary stream data is sent from the sender MPEG-4 application 110 to the sender DMIF 114, it is first processed by the MPEG-4 synchronization layer 140. The packetized, possibly time stamped, possibly multiplexed elementary stream is then transmitted to the receiver DMIF (step 128) and then processed by the receiver MPEG-4 synchronization layer 130 which reassembles the elementary stream as well as ensures timely presentation of the data to the application.

In an alternative embodiment of the present invention, MPEG-2 is equipped with a similar feature. The auxiliary data is inserted and multiplexed at the sending device into an MPEG-2 transport stream. In one preferred embodiment, the auxiliary at the sending device into an MPEG-2 program. This is accomplished by assigning a special stream type value specifying that the type of program element carried within the packets is auxiliary data encoded according to a particular format. Values of stream type have already been defined in

Table 2-29 of the MPEG-2 systems specification (IEC/ISO 13818-1) which permit the announcement of various video, audio and command/control protocol streams. The stream type value assigned to the auxiliary data stream is then 5 transmitted to the receiver by means of the TS program map section() structure defined in Table 2-28. The auxiliary data stream(s) stream type and elementary PID value(s) are added to the relevant program (defined by program number in the TS program map section()) 10 structure). The inner descriptor loop in the TS program map section may be used in connection with the auxiliary data stream type to provide additional, high level, information about the auxiliary data service. The TS program map section() allows the transmission and reconstruction 15 in the receiver of the program map table. The program map table provides the mapping between program numbers and the program elements comprising them. The elementary PID value associated with the auxiliary data stream type specifies the packets conveying the auxiliary 20 data in the MPEG-2 transport stream. The auxiliary data is associated and enhances the service provided by other streams (video or audio elementary streams, for example) published in the program map table. Each program may feature one or several auxiliary data streams. In 25 another embodiment, the auxiliary data streams are announced and associated with other streams like video or audio elementary streams in an MPEG-2 program stream map structure (defined in Table 2-35 of ISO/IEC 13818-1). This is achieved by assigning a particular (stream type, 30 elementary stream ID) tuple to announce a particular auxiliary data stream format. Values of elementary stream ID have already been defined in Table 2-18 of ISO/IEC 13818-1 which can be used to publish the presence of video or audio streams in the multiplexed stream. The 35 inner descriptor loop associated with the auxiliary data stream type may be used to convey additional, high level information about the auxiliary data service. In either

embodiments, and if the program includes reference to at least one video elementary stream, a given auxiliary data stream may be associated with only a portion of the transmitted video portion. Such portion may correspond 5 to a specific region on the television display screen. In other words, the auxiliary data is associated with a particular block of picture elements (pixels) of a scene. Because of encoding mechanism used by MPEG-2 video specification (ISO/IEC 13818-2), it is preferable to 10 define such picture block to coincide with boundaries of a set of contiguous 16x16 pixel blocks. Furthermore, if the data object to which the auxiliary data was associated at the sending device was moving about the screen, the auxiliary data may be reassigned to pixel 15 sets on substantially a frame-to-frame or field-to-field basis.

The terms and expressions which have been employed in the foregoing specification are used therein as terms of description and not of limitation, and there 20 is no intention, in the use of such terms and expressions, of excluding equivalents of the features shown and described or portions thereof, it being recognized that the scope of the invention is defined and limited only by the claims which follow.

WE CLAIM:

1. A video system comprising:
  - (a) a plurality of frames of video each of which is defined by a plurality of scene elements, where said plurality of scene elements for a respective frame together define an image of said respective frame;
  - (b) first auxiliary data descriptive of a first scene element of said plurality of scene elements for said respective frame;
  - (c) second auxiliary data descriptive of a second scene element of said plurality of scene elements for said respective frame; and
  - (d) a sending device that sends said respective frame of said video, including said respective said scene elements, said first auxiliary data, and said second auxiliary data, from said sending device to a receiving device.

2. The video system of claim 1 wherein said video includes audio information.

3. The video system of claim 1 wherein said scene elements are overlapping.

4. The video system of claim 1 wherein said first auxiliary data is a set of descriptors and parameter values resulting from said first scene element.

5. The video system of claim 1 wherein said first auxiliary data is not directly representative of an image.

6. The video system of claim 1 wherein said first auxiliary data is viewable on a display device by selecting said first scene element.

5

7. A video system comprising:

10

- (a) a plurality of frames of video each of which defines an image of said respective frame, each of said frames defined by a plurality of pixels;
- (c) first auxiliary data descriptive of a first portion of said pixels of a respective frame and associated with said first portion of said pixels;
- (c) second auxiliary data descriptive of a second portion of said pixels of said respective frame and associated with said second portion of said pixels, where said first portion is different than said second portion; and
- (d) a sending device that sends said first portion of said pixels, said first auxiliary data, said second auxiliary data, and said second portion of said pixels to a receiving device.

25

8. The video system of claim 7 wherein said video includes audio information.

30

9. The video system of claim 7 wherein said first portion and said second portion are blocks of 16 by 16 pixels of said respective frame.

35

10. The video system of claim 7 wherein said first auxiliary data is a set of descriptors and parameter values resulting from said first portion.

11. The video system of claim 7 wherein said first auxiliary data is not directly representative of an image.

5 12. The video system of claim 7 wherein said first auxiliary data is viewable on a display device by selecting said first portion.

10 13. A video system comprising:

(a) a plurality of frames of video each of which is defined by a plurality of scene elements, where said plurality of scene elements for a respective frame together define an image of said respective frame;

15 (b) each of said scene elements defined by a scene description format including a scene element data object format that defines the number of elementary streams and elementary stream descriptors;

20 (c) each of said elementary stream descriptors including a decoder type identifier;

(d) each of said decoder type identifiers represents at least one of an audio stream, object descriptor, and auxiliary data, with at least one of said decoder type identifiers represents said auxiliary data; and

25 (e) said auxiliary data descriptive of a respective one of said scene elements.

30 14. A system for encoding a moving picture comprised of a sequence of time separated frames depicting scenes, said system comprising a scene description format including a scene element data object format that includes an elementary stream descriptor format including, in turn, a decoder type identifier format wherein any one of a set of decoder types may be

represented by one out of a set of predetermined bit values and one of said predetermined bit values represents an auxiliary information decoder type.

5               15. The system of claim 14 wherein an audio stream is also encoded.

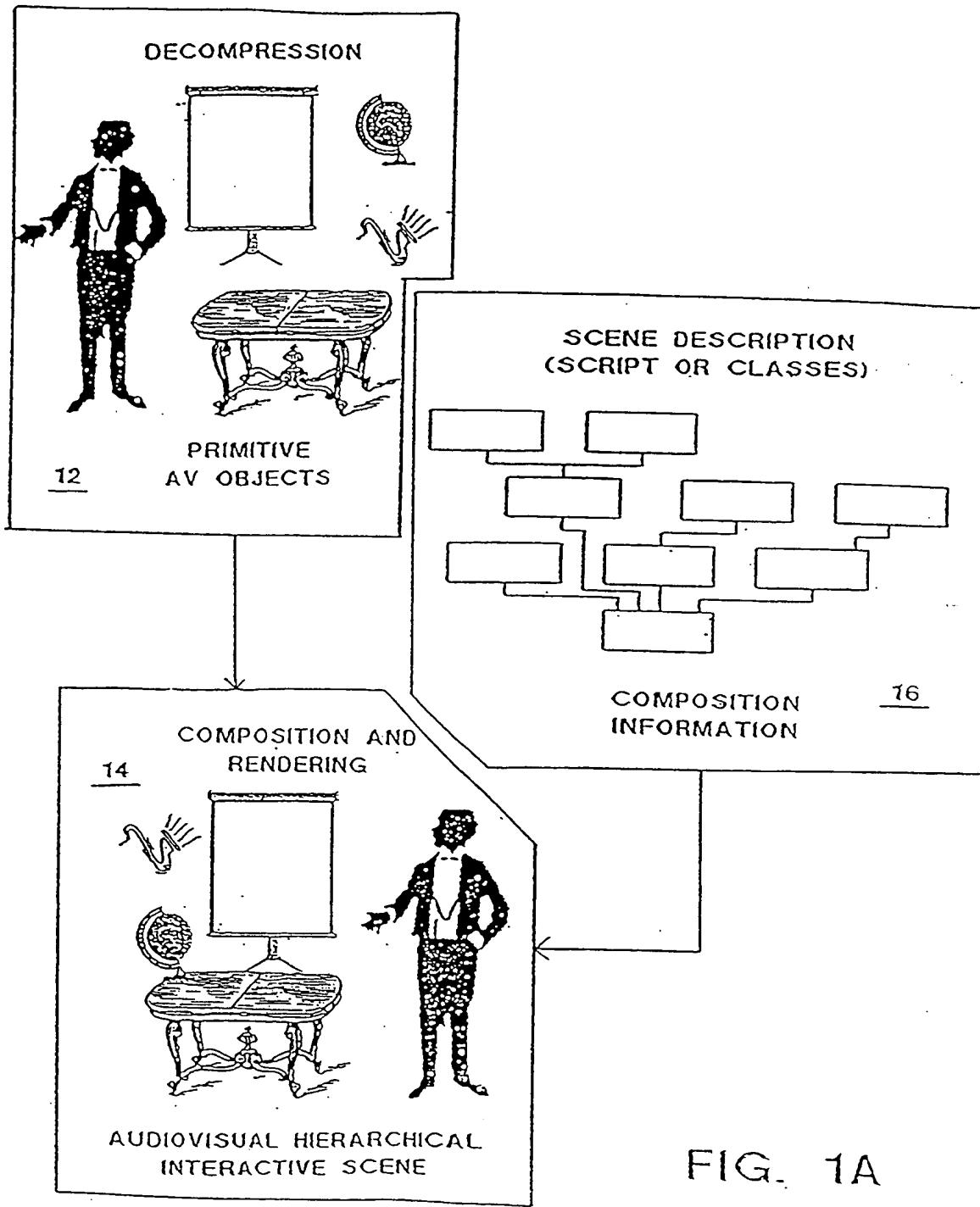
10              16. A system for encoding a moving picture comprised of a sequence of time separated frames, said system divides at least one of the encoded frames into a plurality of pixel groups and further including a system for sending auxiliary data, and associating said auxiliary data to at least one pixel group.

15              17. The system of claim 16 wherein an audio stream is also encoded.

20              18. A system for encoding a moving picture comprised of a sequence of time separated frames, said system adapted to divide said frames into scene elements and including a receiver display subsystem adapted to receive and display said frames and a pipe assignment subsystem adapted assign pipes to the scene elements and wherein said subsystem assigns at least one of a logical 25 and physical transmission channel designator to each pipe and transmits a quantity representative of said pipe designator to said receiver display subsystem.

30              19. The system of claim 18 wherein an audio stream is also encoded.

1/5



SUBSTITUTE SHEET (RULE 26)

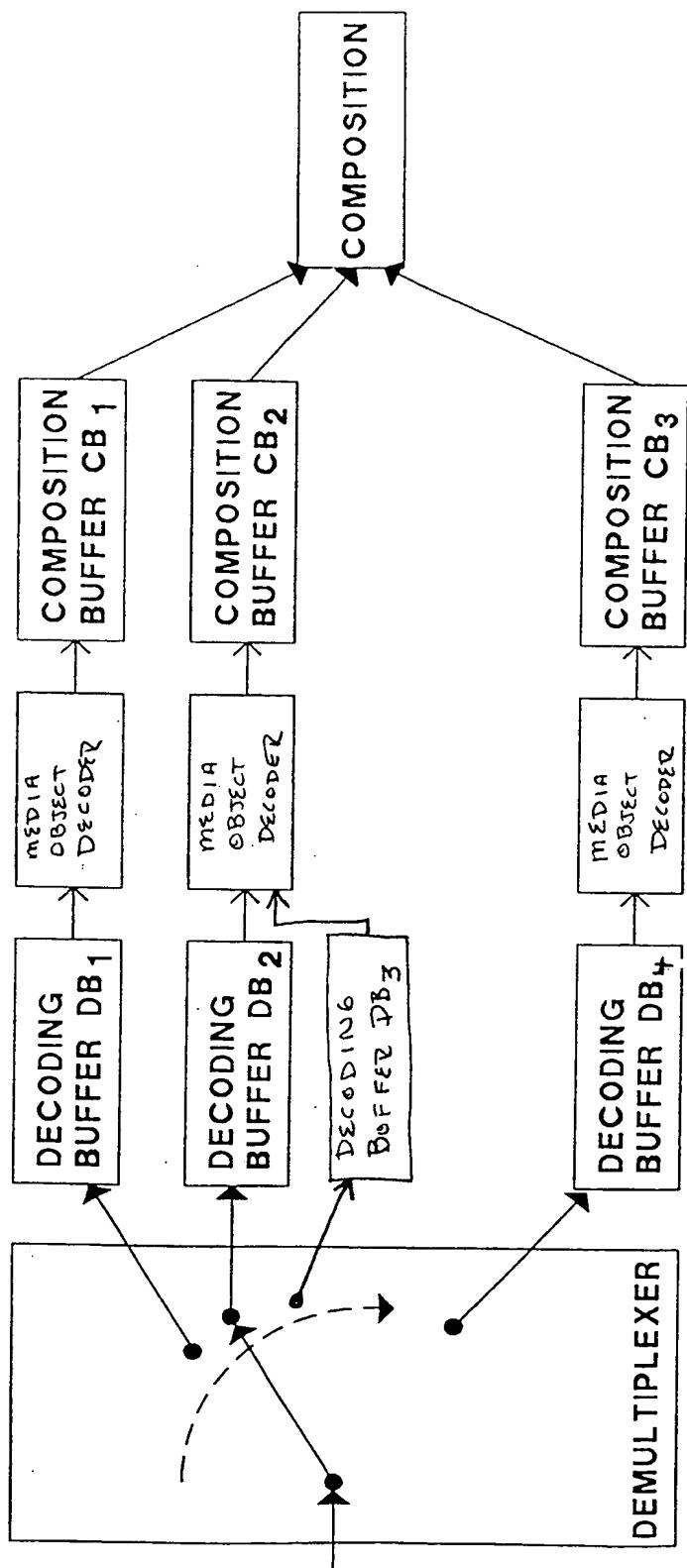


FIG. 1B

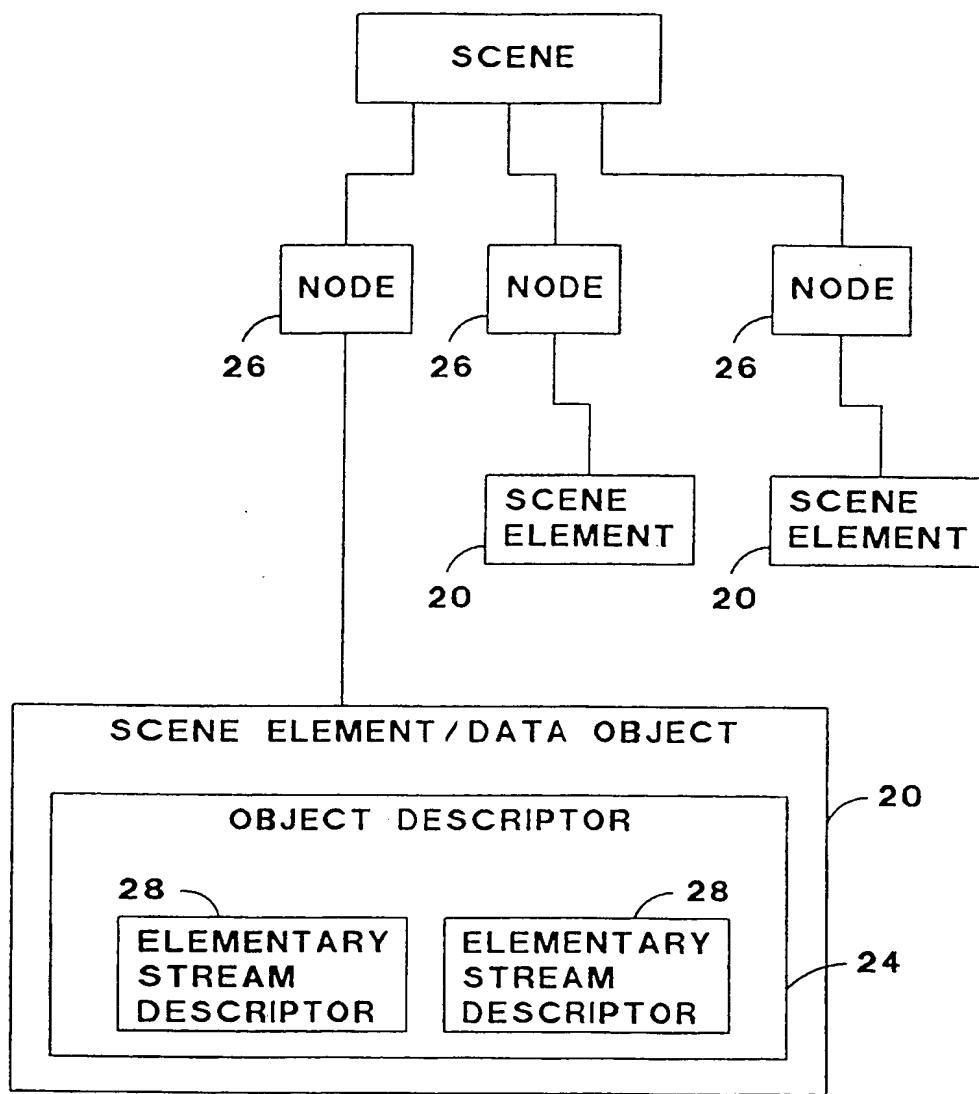


FIG. 2

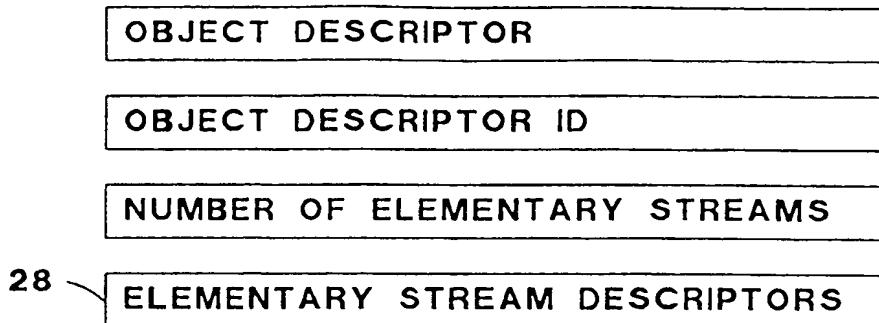


FIG. 3

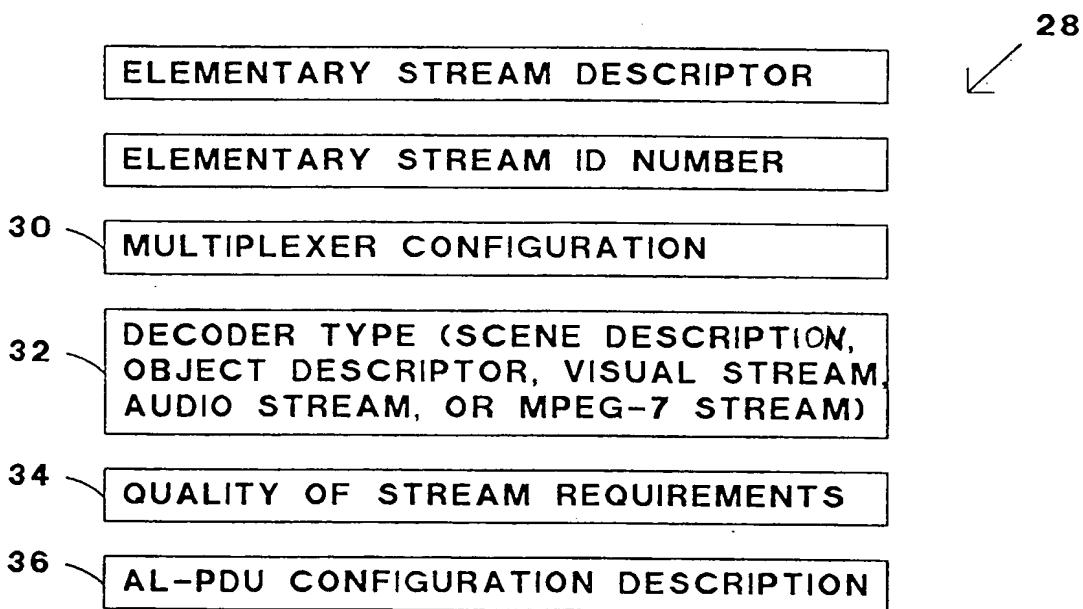


FIG. 4

5/5

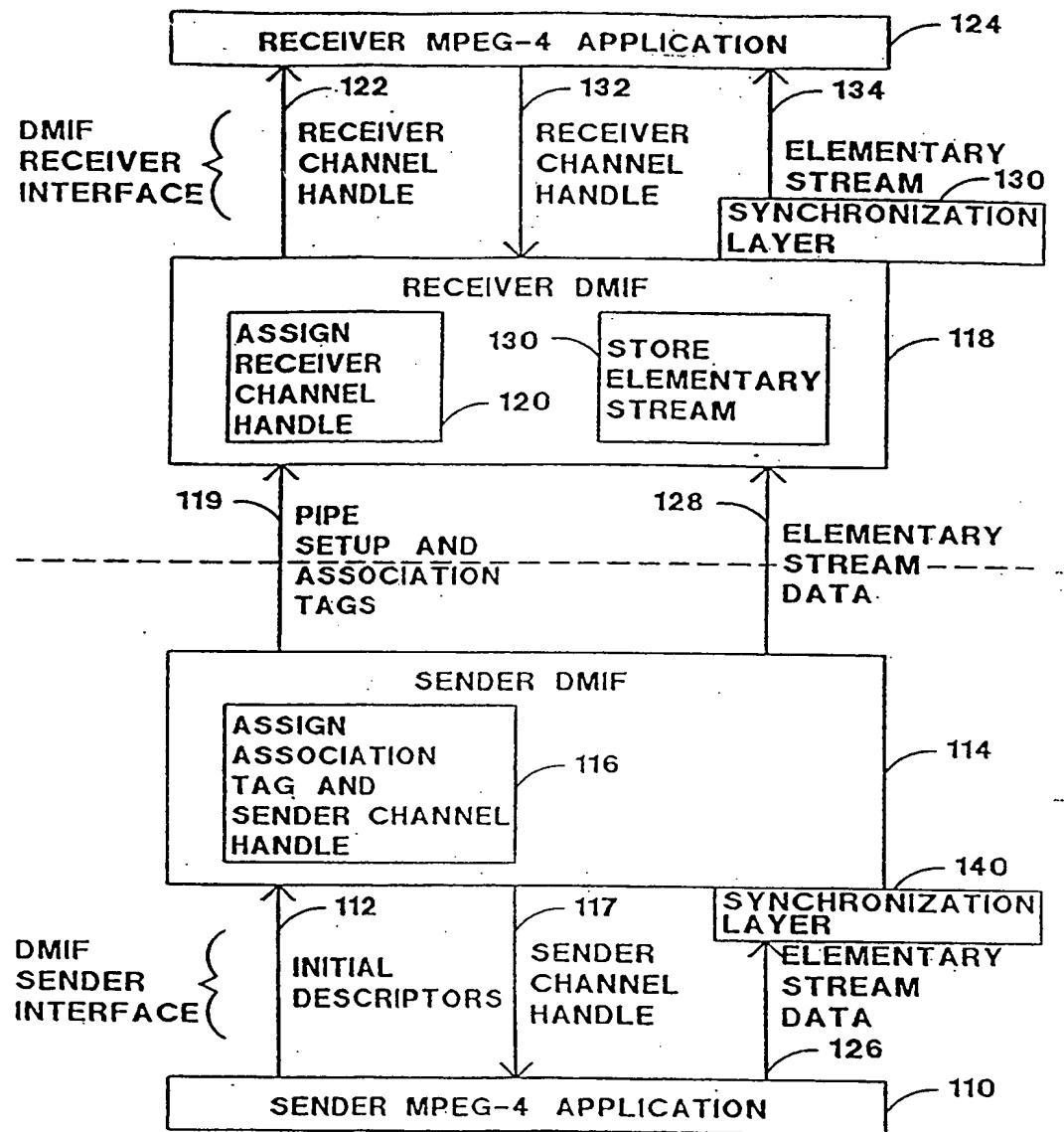


FIG. 5

SUBSTITUTE SHEET (RULE 26)

## INTERNATIONAL SEARCH REPORT

.onal Application No

PCT/IB 98/01042

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 H04N7/26 H04N7/52 G06F17/30

According to International Patent Classification(IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 H04N G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOENEN R ET AL: "MPEG-4: Context and objectives" SIGNAL PROCESSING. IMAGE COMMUNICATION, vol. 9, no. 4, May 1997, page 295-304 XP004075332 see page 297, paragraph 3.1 - page 298, line 21 see page 300, paragraph 3.3; figure 1 ---	1-19
A	EBRAHIMI T: "MPEG-4 VIDEO VERIFICATION MODEL: A VIDEO ENCODING/DECODING ALGORITHM BASED ON CONTENT REPRESENTATION" SIGNAL PROCESSING. IMAGE COMMUNICATION, vol. 9, no. 4, May 1997, pages 367-384, XP000700946 see page 370, paragraph 2 - page 373, paragraph 2.1; figures 1-3 ---	1-19

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier document but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  
 "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 October 1998	14/10/1998
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	Authorized officer  Foglia, P

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 98/01042

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BURRILL V ET AL: "TIME-VARYING SENSITIVE REGIONS IN DYNAMIC MULTIMEDIA OBJECTS: A PRAGMATIC APPROACH TO CONTENT BASED RETRIEVAL FROM VIDEO" INFORMATION AND SOFTWARE TECHNOLOGY, vol. 36, no. 4, 1 January 1994, pages 213-223, XP000572844 see page 215, left-hand column, line 1 - right-hand column, line 1 ---	1-12, 16, 17
X	WO 97 12342 A (WISTENDAHL DOUGLASS A ;CHONG LEIGHTON K (US)) 3 April 1997 see abstract see page 3, line 11 - line 35 see page 18, line 14 - page 19, line 39; figures ---	1-12, 16, 17
A	"MULTIMEDIA HYPERVIDEO LINKS FOR FULL MOTION VIDEOS" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 37, no. 4A, 1 April 1994, page 95 XP000446196 see page 95 ---	1-12, 16, 17
A	EP 0 596 823 A (IBM) 11 May 1994  see abstract see column 2, line 1 - line 45; claims -----	1-12, 16, 17

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 98/01042

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9712342 A	03-04-1997	US	5708845 A	13-01-1998
EP 0596823 A	11-05-1994	US	5539871 A	23-07-1996
		JP	2677754 B	17-11-1997
		JP	7085243 A	31-03-1995

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

This Page Blank (uspto)